

Hofer, Manfred; Köpke, Bernd

Die Kategorisierung von Schülern durch Lehrer: eine Methodenstudie

Unterrichtswissenschaft 15 (1987) 3, S. 285-294



Quellenangabe/ Reference:

Hofer, Manfred; Köpke, Bernd: Die Kategorisierung von Schülern durch Lehrer: eine Methodenstudie
- In: Unterrichtswissenschaft 15 (1987) 3, S. 285-294 - URN: urn:nbn:de:0111-pedocs-54544 - DOI:
10.25656/01:5454

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-54544>

<https://doi.org/10.25656/01:5454>

in Kooperation mit / in cooperation with:

BELTZ JUVENTA

<http://www.juventa.de>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Digitalisiert

Mitglied der


Leibniz-Gemeinschaft

Unterrichts- Wissenschaft

Lernen in Schule,
Beruf und Freizeit

Heft 3 – 1987

Mitteilung an unsere Leser und Autoren 258

Thema: Denkprozesse von Lehrern

Verantwortlicher Herausgeber für den Thementeil:
Prof. Dr. Manfred Hofer

Manfred Hofer: Einführung 259

Hartmut-A. Oldenburger: Lehrerkognitionen über Schülereigenschaften – Theoretische und methodologische Perspektiven 261

Bernhard Kraak: Was Lehrerinnen und Lehrer denken und tun, erklärt mit der Handlungs-Entscheidungs-Theorie 274

Manfred Hofer, Bernd Köpke: Die Kategorisierung von Schülern durch Lehrer: eine Methodenstudie 285

David C. Berliner: Der Experte im Lehrerberuf: Forschungsstrategien und Ergebnisse 295

Hanns-Dietrich Dann, Kurt-Christian Tennstädt, Winfried Humpert, Frank Krause: Subjektive Theorien und erfolgreiches Handeln von Lehrern/-innen bei Unterrichtskonflikten 306

Gerhard Kaminski: Einige Reflexionen über Forscher-Reflexionen über Lehrer-Reflexionen. Eine Buchbesprechung 321

Allgemeiner Teil

Karl H. Reich: Religiöse und naturwissenschaftliche Weltbilder: Entwicklung einer komplementären Betrachtungsweise in der Adoleszenz 332

Buchbesprechungen 344

Die Themen der nächsten Hefte:

4/1987 Schreiben als Forschungsgegenstand (hrsg. von G. Eigler)

1/1988 Interkulturelles Lernen (hrsg. von H.-J. Krumm)

2/1988 Lernorte und Lernprozesse in der beruflichen Bildung (hrsg. von F. Achtenhagen)

3/1988 Medien und Anschauung (hrsg. von P. Strittmatter)

Die Kategorisierung von Schülern durch Lehrer: eine Methodenstudie

Ausgangspunkt dieser Untersuchung ist die Kategorisierungstheorie, in der angenommen wird, daß Lehrer Schüler nach einem erschöpfenden und disjunkten Kategoriensystem gruppieren: ein konkreter Schüler wird durch Merkmalsvergleich in jene kognitive Kategorie eingeordnet, zu deren Prototypen er die größte Ähnlichkeit besitzt. In der Untersuchung wurde überprüft, wie sehr die Ergebnisse zweier Erhebungsmethoden übereinstimmen, wie groß die intraindividuelle Konstanz der Ergebnisse und wie stark deren interindividuelle Variation ist. Als Methoden wurden die freie Gruppierungsaufgabe und die Beurteilung von Schülern nach Merkmalen mit anschließender Clusteranalyse gewählt. 37 Lehrer sollten beide Methoden bei den Schülern zweier Schulklassen, die sie unterrichten, anwenden. Die Untersuchung war so angelegt, daß jede Klasse von zwei verschiedenen Lehrern beurteilt wurde. Beim Vergleich der beiden Methoden im Hinblick auf gleiche Schulklassen konnten nur schwache Korrespondenzen festgestellt werden. In verschiedenen Klassen kamen die gleichen Lehrer nur zum Teil zu ähnlichen Gruppierungen. Die Gruppierungen verschiedener Lehrer in bezug auf dieselbe Klasse zeigten die geringsten Übereinstimmungen. Es wird vermutet, daß das Modell der Punktwolken im euklidischen Raum nicht die geeignete Abbildungsform zur Darstellung von Kognitionen von Lehrern über Schüler ist.

How teachers categorize students

Categorization theory suggests that teachers cognitively associate students on a set of prototype categories. This study investigates three questions about this theory: 1) does evidence of categorization depend on how teachers' notions of categories are collected? 2) is there intra-individual consistency of categories? and 3) are there inter-individual variations in prototypes? Two methods of data collection were used, evaluation of students on polar adjectives with cluster analysis of the scores for each teacher, and on a pupil sort task. Each of 36 teachers from various levels of schooling applied these methods to the students from two different classes. The investigation included two teachers for every class to control for class specific variation. Evidence of categorization was found with both types of data collection, but the correspondence between the two methods was weak. There was not much evidence of intra-individual consistency, in that teachers tended to group students of different classes differently. Also there was no evidence of inter-teacher consistency in the use of prototypes. It is suggested that a Euclidian model of categorization across a three-dimensional space does not reflect the way teachers actually think about their students.

1. Problemstellung

Seit den Untersuchungen von *Brophy* und *Good* (1974) zum Lehrerverhalten gegenüber verschiedenen Schülern, wurde der Frage Aufmerksamkeit geschenkt, nach welchen Kriterien Schüler zu unterscheiden oder auszuwählen seien. Immer dann, wenn das Lehrerverhalten nicht in bezug auf die ganze Klasse, sondern in bezug auf einzelne Schüler Gegenstand der Analyse ist, stellt sich das Problem, Kriterien festzulegen, nach denen die Schüler ausgesucht oder eingeteilt werden. Dabei sind grundsätzlich zwei Möglichkeiten denkbar. Das Kriterium kann nach der speziellen Fragestellung des Forschers definiert werden. So mag das Interesse einer wissenschaftlichen Fragestellung darauf gerichtet sein, ob Lehrer Jungen und Mäd-

chen in bestimmten Bereichen unterschiedlich behandeln (z.B. *McFarland* 1969). Dagegen lassen sich Verfahren abheben, die sich anstatt an äußeren Kriterien daran orientieren, welche eigenen Merkmale Lehrer bei der Beurteilung von Schülern anlegen, bzw. nach welchen Merkmalen sie die Schüler ihrer Klasse vorwiegend unterscheiden.

Einige Untersucher wählten ein Vorgehen, das eine Mischung aus beiden Verfahren darstellt. So ließ *Silberman* (1969) von Lehrern jene Schüler auswählen, zu denen sie eine bestimmte vorgegebene Beziehung hatten. *Garner* und *Bing* (1973) sowie *Cunningham* (1975) suchten mit Hilfe statistischer Techniken (automatische Klassifikation und Q-Analyse) nach Gruppierungen, wobei sie für jeden Schüler Beurteilungsdaten der Lehrer erhoben, aber auch objektive Angaben (wie Prüfungsergebnisse) verwendeten.

Gilt das Interesse Denkprozessen von Lehrern und der Frage, wie das Denken und Verhalten organisiert ist, so liegt es nahe, bei der Suche nach Einteilungskriterien möglichst wenig externe Vorgaben zu machen und die von den Lehrern selbst als für sie relevant erachteten Kriterien zu verwenden. In diesem Fall erhebt sich das Problem, die subjektiven Kriterien der Lehrer methodisch einwandfrei und theoretisch angemessen zu ermitteln. *Thelen* (1967) war der erste, der den Versuch unternommen hat, Schülertypen aus der Sicht des Lehrers empirisch zu erheben. Er bat Gymnasiallehrer, selbsterlebte Schülertypen frei zu beschreiben. Die erhaltenen Charakterisierungen konnte er zu 26 Einzeltypen zusammenfassen. Darunter fallen Bezeichnungen wie „Träumer“, „Lehrer-Beeindrucker“, „Schönheitskönigin“, „Leidender“, „Athlet“. *Thelen* (1967, S. 65 ff.) nahm eine weitere Reduktion auf vier Haupttypen vor: gute Schüler, gleichgültige Schüler, schlechte Schüler, „verlorene Seelen“. Ein anderer Ansatz ist der „pupil sort task“, wie ihn *Morine-Dershimer* (1978) durchführte. Die Namen der Schüler wurden auf je eine Karte geschrieben und den Lehrern mit der Bitte vorgelegt, sie zu ordnen. Die Karten mit den Namen jener Schüler, die sich ähnlich verhalten, sollten zusammengelegt werden. Anschließend sollten die Lehrer die Schüler jeder Gruppe nach ihren Gemeinsamkeiten beschreiben und die Unterschiede zwischen den Gruppen herausheben. Eine weitere Technik wendete *Hofer* (1981; 1986) an. Lehrer wurden gebeten, alle Schüler ihrer Klasse hinsichtlich Eigenschaftsdimensionen zu beurteilen. Für die Schüler wurde eine automatische Klassifikation auf der Basis der Lehrerbeurteilungen durchgeführt. Die ermittelten Gruppierungen ließen sich anhand ihres jeweiligen Eigenschaftsprofils gut beschreiben. Die Ergebnisse dienten als Ausgangspunkt für eine weitere Untersuchung. Darin zeigte sich, daß Lehrer sich in ihrem Verhalten gegenüber bestimmten Schülern, die als typische Vertreter von vier Gruppen ausgewählt worden waren (von: „Klassenprimus“, „schlechter Schüler“, „extravertierter Schüler“ und „introvertiert-sensibler Schüler“), systematisch unterschieden. Hier wurde davon ausgegangen, daß Lehrer ihr Wissen über Schüler in einem Schüler-Kategoriensystem ordnen. Ein Kategoriensystem ist eine Menge von Klassen, innerhalb derer Objekte nach ihrer Ähnlichkeit in Merkmalen als äquivalent angesehen werden. Eine Kategorie kann durch Angabe ihres klarsten Vertreters, des Prototypen, angegeben werden. Nach *Rosch* (1978) sowie

Cantor und Mischel (1979) hat ein Kategoriensystem von mittlerem Abstraktionsgrad den größten Nutzen für die Person. Es maximiert Vielfalt, Differenzierung und Lebendigkeit, ohne daß eine übergroße Differenzierung zur Last wird. Hofer (1986) nimmt an, daß Lehrer über eine Reihe von Handlungsprogrammen verfügen, die mit bestimmten Kategorien über Schüler assoziiert werden und für diese spezifisch sind. Wird von einem konkreten Schüler eine bestimmte Schülerkategorie (z.B. Klassenprimus) instantiiert, so führt dies zu einer Aktivierung des zugehörigen (passenden) Handlungsentwurfs. Zur empirischen Ermittlung der subjektiven Schülerkategorien werden Methoden wie die o.g. Schülersortieraufgabe und die Clusteranalyse von Beurteilungsskalen herangezogen. Die von den Lehrern zu den Gruppen abgegebenen Benennungen bzw. die charakteristischen Merkmalsprofile werden als Repräsentanten der jeweiligen Prototypen betrachtet.

2. Fragestellung

Eine Konzeptualisierung der handlungssteuernden Funktion von subjektiven Schülerkategorisierungen wirft eine Reihe von miteinander verbundenen Problemen auf, die vor allem methodischer und theoretischer Art sind. Drei davon geben Anlaß für die vorliegende Untersuchung.

- (a) Führen die Methoden der Schülersortierung und der Clusteranalyse von Beurteilungsskalen zu vergleichbaren Ergebnissen? Beide Methoden liefern eine Aufteilung der Schüler einer Klasse in verschiedene Gruppen sowie eine Charakterisierung jeder Gruppe.
- (b) Ist die für einen Lehrer ermittelte Kategorisierung invariant gegenüber den verwendeten Reizen? Resultieren die gleichen Ergebnisse, wenn derselbe Lehrer das Verfahren bei Schülern aus zwei verschiedenen, ihm bekannten Schulklassen durchführt? Die Kategorien sollen Kognitionen der Lehrer bezeichnen, die unabhängig von konkreten Schülern aufgefaßt werden.
- (c) Verfügen verschiedene Lehrer über vergleichbare Kategoriensysteme? In der Arbeit von Hofer (1981) wurde die Clusteranalyse über alle Lehrer der untersuchten Stichprobe gemeinsam durchgeführt. Analysen hatten gezeigt, daß die Ergebnisse der Clusteranalysen einzelner Schulklassen in circa 70% der Fälle mit dem Ergebnis der Gesamtclusteranalyse übereinstimmten. Auch Morine-Dershimer (1978) beobachtete, daß bei einem Großteil der freien Gruppierungen (80–90%) Gemeinsamkeiten zwischen Lehrern festzustellen waren. Die Frage nach dem Grad der Gemeinsamkeit von Schülerkategorien kann zweckmäßigerweise geprüft werden, wenn man verschiedene Lehrer, die in gleichen Klassen unterrichten, in die Untersuchung einbezieht. Auf diese Weise können Unterschiede in der Zusammensetzung der Schulklasse konstant gehalten werden.

Zusammenfassend können die Nullhypothesen dieser Untersuchung wie folgt formuliert werden:

- Die beiden Methoden „Freie Schülergruppierung“ und „Clusteranalyse von Persönlichkeitsbeurteilungen“ führen, wenn sie bei den gleichen Lehrern angewendet werden, zu vergleichbaren Ergebnissen.
- Die Ergebnisse, die von denselben Lehrern bezogen auf verschiedene Schulklassen erzielt werden, sind vergleichbar.
- Die Ergebnisse, die von verschiedenen Lehrern bezogen auf die gleiche Schulklasse erzielt werden, sind vergleichbar.

3. Methode

3.1 Durchführung

37 Gymnasial- und Realschullehrer aus fünf Schulen wurden gebeten, für zwei Klassen der Jahrgangsstufen 5 bis 11, in denen sie unterrichteten, die Schüler mit Hilfe von Namenskartchen pro Klasse so zusammenzufassen, daß ähnliche Schüler eine Gruppe bilden. Jede Schülergruppierung sollte anschließend mit eigenen Worten beschrieben werden. Die Aufgabe schloß mit der Aufforderung, jeder Gruppe einen Namen zu geben. Die Auswahl der Klassen wurde vom jeweiligen Schulleiter vorgenommen. Es wurde darauf geachtet, daß die Variablen „Dauer der Bekanntschaft“ und „Klassenstufe“ nach Möglichkeit konstant waren.

Im zweiten Teil der Erhebung sollten dieselben Schüler einzeln beurteilt werden. Dazu wurden den Lehrern Rating-Bögen mit 25 bipolaren Merkmalen auf 7stufigen Skalen, ähnlich wie bei Hofer (1981), vorgegeben. Die Beurteilungsbögen wurden für jeden Lehrer und jede Klasse mit einer hierarchisch-agglomerativen (Ward 1963) und einer iterativ-relokativen (Wishart 1969) Clusteranalyse ausgewertet. Die Gruppenzahl wurde durch das für hierarchisch-agglomerative Clusteranalyse-Verfahren anwendbare Varianz-Kriterium ermittelt.

Fünf der 37 Lehrer gaben die Fragebögen nicht oder unbearbeitet zurück. Von den restlichen 32 Lehrern gruppieren und beurteilten sechs nur die Schüler einer Klasse.

3.2 Auswertung

Auf der Grundlage der erhobenen Gruppierungen und Gruppenbeschreibungen wurden die folgenden Auswertungen vorgenommen:

- (a) *Methodenvergleich.* Zur Beantwortung der Frage nach methodenbedingten Einflüssen auf die Abbildung subjektiver Schülergruppierungen wurde lehrer- und klassenweise
 - die Gruppenzahl bei freier Gruppierung mit der nach dem Varianzkriterium optimalen Clusterzahl verglichen, und
 - die Übereinstimmung der Schülerzuordnungen von freien Gruppierungen mit numerischen Klassifikationen bestimmt; für diese Analyse wurde die Gruppenzahl gewählt, die Lehrer bei freier Klassifikation generierten.
 - (b) *Intraindividueller Vergleich.* Um festzustellen, welche situativen und klassenspezifischen Einflüsse bei der Gruppierung von Schülern auftreten, verglichen wir die Kategorisierungen, die Lehrer für zwei verschiedene Klassen vorgenommen hatten.
 - (c) *Interindividueller Vergleich.* Zur Ermittlung lehrerspezifischer Kategorisierungen wurden für identische Schulklassen die Schülergruppierungen verschiedener Lehrer miteinander verglichen.
- Die beiden Vergleiche (b) und (c) wurden mit jeder der beiden Methoden unter (a) durchgeführt. Dies kann als eine Art von Replikation aufgefaßt werden.

3.3 Koeffizienten

Die methoden- und personenspezifischen Gruppenunterschiede wurden mit Hilfe der folgenden Koeffizienten bestimmt:

(1) *Tanimoto-Koeffizient:*

Der Tanimoto-Koeffizient s (Bergs 1980, S. 151f.) ist das arithmetische Mittel aus der Summe der Übereinstimmungsquotienten $t(ij)$ bei n Schnittmengen. Die Indizes i und j stehen für die jeweilige Partition einer Klassifikation. Die $t(ij)$ -Quotienten und damit auch der Tanimoto-Koeffizient können Werte zwischen 0 und +1 annehmen. Ein s -Koeffizient von +1 bedeutet, daß beide Klassifikationen völlig identisch sind. Per Simulation wurde der Bereich von zufälligen Übereinstimmungen ermittelt. Dieser Zufallsbereich reicht je nach Größe der Schulklasse von 0.15 bis 0.24.

(2) *Schnittmengen-Koeffizient:*

Der Tanimoto-Koeffizient vernachlässigt es, Gruppen einzubeziehen, die sich nicht überschneiden. Er unterschätzt die tatsächliche Ähnlichkeit. Aus diesem Grunde wurde als Ergänzung ein zweiter Koeffizient entwickelt. Der Gruppenähnlichkeits-Koeffizient k mißt die Überlappungen der Gruppen und setzt sie in Relation zu der kleinst- und größtmöglichen Überschneidung. Die Ähnlichkeit von Gruppierungen ist umso größer, je weniger Partitionsüberschneidungen zwischen den einzelnen Klassifikationen auftreten. Der k -Koeffizient errechnet sich wie folgt: $k = 1 - (v-g)/(u-g)$, wobei u die Zahl der maximal möglichen Überschneidungen, v die Zahl der tatsächlichen Überschneidungen und g die größte vorkommende Gruppenzahl und damit die kleinste Überschneidungsmöglichkeit bezeichnet. Durch g und u ist der Wertebereich von Gruppenüberschneidungen definiert. Der Überschneidungs-Koeffizient k kann wie der Tanimoto-Koeffizient s einen Wert zwischen 0 und +1 annehmen. Auch hier bedeutet ein Wert von +1 völlige Identität. Der Erwartungswert des k -Koeffizienten liegt höher als der beim Tanimoto-Koeffizienten. Je nach Klassengröße kann er bis 0,4 reichen.

4. Ergebnisse

4.1 Methodenvergleich

4.1.1 Freie Gruppierung

Die Schulklassen, die den Schülergruppierungen zugrunde lagen, variierten in ihren Klassenstärken zwischen 15 und 34 Schülern. Die 32 Lehrer bildeten in 44 verschiedenen Klassen im Mittel sechs Gruppen. Obwohl den Lehrern vorgegeben war, möglichst homogene (in sich geschlossene) Gruppen zu bilden, generierten sechs der 32 Lehrer heterogene Gruppen, indem sie Mehrfachzuordnungen vornahmen. Bei den 37 homogenen Kategorisierungen ergab sich zwischen Gruppenzahl und Klassengröße eine Korrelation von 0,05.

Die Beschreibungen, mit denen die Lehrer die von ihnen gebildeten Gruppen charakterisierten, wurden in acht Merkmalskategorien angeordnet, um sie inhaltlich vergleichen zu können. In der Reihenfolge der Häufigkeit ihres Auftretens: Persönlichkeitseigenschaften, Anstrengung/Mitarbeit, Begabung/Leistung, Aktivität/Dynamik, Diszipliniertheit, Sozialverhalten, seelische Robustheit, Reife. Der abschließenden Aufgabe zur Etikettierung jeder gebildeten Gruppe kamen 24 Lehrer nach. Auf eine quantitative Auswertung wurde verzichtet. Einige Beispiele mögen den Reichtum an Idiosynkrasien veranschaulichen: Kindsköpfe, Klassenkern, Schleimer, Gänschen, Fans, Schlitzohren, Mitläufer, Problemgruppe, brave Mäuschen, Kasper, blinde Hühner, stille Wasser, Lolitas. Eine intuitive Zuordnung dieser Etiketten zu den vier Typen nach Hofer (1981) und Thelen (1967) konnte anhand der genannten Eigenschaften in durchschnittlich 82% aller Fälle erreicht werden.

4.1.2 Clustergruppierung

Deutlich erkennbare Clusterlösungen ergaben sich nach dem Kriterium des Fehleranstiegs nur bei knapp der Hälfte der Analysen. Dies läßt den Schluß zu, daß bei der anderen Hälfte die untersuchten Objektmenge(n) entweder gar keine Cluster enthalten oder aus solchen Clustern bestehen, die das Verfahren nicht erkennen kann. Wir haben bei schwer erkennbaren Elbows das schwächere Tangentenkriterium zur Clusterfindung herangezogen. Nach dem hierarchisch-agglomerativen WARD-Verfahren resultierten im Mittel 3,5 Cluster. Dreier- und Viererlösungen kamen mit 30 und 35 Prozent am häufigsten vor. Die Bandbreite der Gruppenzahl reichte von zwei (19 Prozent) bis zu sechs (3 Prozent) Clustern. Im Mittel wurden einem Cluster 6,5 Schüler zugeordnet.

4.1.3 Vergleich der Methoden

(a) *Gruppenzahl.* Bildet man für jeden Lehrer klassenweise die Differenz zwischen der Gruppenzahl bei freier Gruppierung und der Clusterzahl, erhält man eine durchschnittliche Abweichung von 2,5 Gruppen. In zwei von 36 Fällen wurden bei der Clusteranalyse mehr Gruppen generiert als bei freier Gruppierung. In einem Drittel der Fälle war die Differenz zwischen frei gebildeten Gruppen und Clustern größer als drei. Die größte Gruppendifferenz lag bei sieben.

(b) *Übereinstimmung der gruppierten Schüler.* Die Verteilung der Tanimoto- ($s = 0.35$) und Schnittmengen-Koeffizienten ($k = 0.62$) für die iterativ-relokative Lösung zeigt, daß die Übereinstimmung der Zuordnungen nicht sehr hoch ist. Doch liegen die Übereinstimmungen deutlich höher als durch Zufall zu erwarten.

4.2 Intraindividueller Vergleich

4.2.1 Freie Gruppierung

(a) *Gruppenzahl.* Dieselben Lehrer bildeten in verschiedenen Schulklassen meist unterschiedlich viele Gruppen. Die Gruppenzahl differierte durchschnittlich um 1,4 Einheiten. Rund 81 % der Gruppierungen wichen ein bis zwei Gruppen voneinander ab.

(b) *Merkmalsvergleich.* Der Vergleich der Gruppen kann nicht auf der Basis derselben Elemente vorgenommen werden. Verglichen werden müssen für jeden Lehrer Ergebnisse, die an verschiedenen Schülern ermittelt worden sind. Grundlage für den Vergleich waren die von den Lehrern vorgenommenen Beschreibungen der gebildeten Gruppen. Zunächst wurde pro Beschreibung inhaltsanalytisch festgestellt, wieviele Beschreibungselemente jeder Lehrer bei seiner Charakterisierung verwendete. Die Lehrer gebrauchten durchschnittlich 3,3 Beschreibungselemente. Jedes benutzte Beschreibungselement wurde nach den genannten acht Merkmalen inhaltlich klassifiziert. Es wurde nun für jedes der acht Merkmale pro Lehrer festgestellt, ob es in beiden Schulklassen Verwendung fand, nur in der Schulklasse A oder nur in der Schulklasse B. Merkmalsidentität war in 64 % aller Fälle feststellbar. In allen anderen Fällen wurde ein Beschreibungsmerkmal nur in einer der beiden gruppierten Klassen verwendet. Besonders hohe Übereinstimmungen ergaben sich hinsichtlich der Merkmale Anstrengung/Mitarbeit, Aktivität/Dynamik und seelische Robustheit, niedrige bei Reife und Disziplinertheit. Nur bei 16 % der Lehrer, die in den intraindividuellen Vergleich einbezogen werden konnten, war eine 100prozentige Merkmalsübereinstimmung festzustellen.

4.2.2 Clustergruppierung

(a) *Gruppenzahl.* Bei numerischer Gruppierung nach dem hierarchisch-agglomerativen WARD-Verfahren ergab sich ein mittlerer Unterschied in der Gruppenzahl von 1,3 Gruppen. Wenn man bedenkt, daß die mittlere Clusterzahl 3,5 betrug, so ist diese Differenz recht erheblich.

(b) *Merkmalsvergleich.* Der Vergleich von Gruppierungen, die Lehrer in verschiedenen Klassen vorgenommen haben, ist beim Verfahren, das mit skalierten Beurteilungen arbeitet, durch Korrelation von Clustermerkmalsprofilen möglich. Bei jenen Lehrern, für die in beiden Klassen gleich viele Cluster resultierten, wurden die Profile einer Klasse mit denen der anderen Schulklasse korreliert. Die Kriterien, die für übereinstimmende Clusterbildungen formuliert worden waren (hohe Korrelationen zwischen je zwei Clustern, niedrige Korrelationen sonst), wurden nicht erfüllt.

4.3 Interindividueller Vergleich

4.3.1 Freie Gruppierung

(a) *Gruppenzahl.* Bei freier (homogener) Gruppierung differierte die Gruppenzahl im Mittel um zwei Einheiten; die maximale Abweichung betrug sechs Gruppen. Nur 20 % der klassenbezogenen Gruppierungen waren zahlenmäßig identisch, 27 % wichen um eine, 33 % um zwei Gruppen ab. Um mehr als zwei Gruppen differierten 20 %. Da es sich jeweils um dieselben Schüler handelte, müssen die festgestellten Unterschiede auf Differenzen in den Lehrersichtweisen zurückgeführt werden.

(b) *Übereinstimmung der gruppierten Schüler.* Der durchschnittliche Tanimoto-Koeffizient betrug $s = 0.24$, der mittlere k -Koeffizient 0.48 . Damit liegen sie sehr nahe an den maximalen Erwartungswerten.

(c) *Merkmalsvergleich.* Die von den Lehrern vorgenommenen Beschreibungen der gebildeten Gruppen wurden miteinander verglichen. Nur in 48 % der Fälle verwendeten die Lehrer die gleichen Merkmale, um die von ihnen gebildeten Gruppen von Schülern der gleichen Klasse zu beschreiben.

4.3.2 Cluster-Gruppierung

(a) *Gruppenzahl.* Bei numerischer Gruppierung nach dem hierarchisch-agglomerativen WARD-Verfahren ergab sich eine durchschnittliche Abweichung von einem Cluster und eine maximale Differenz von drei Gruppen. Die gleiche Gruppenzahl trat nur in 25 % der Fälle auf.

(b) *Übereinstimmung der gruppierten Schüler.* Die numerischen Übereinstimmungen waren bei der Cluster-Gruppierung fast gleich hoch wie bei der freien Gruppierung ($s = 0.25$; $k = 0.46$). Etwas höher lagen die Werte, bildet man den Durchschnitt nur für Partitionen mit übereinstimmender Gruppenzahl.

(c) *Merkmalsvergleich.* Die Korrelationsmatrizen zwischen Clustermerkmalsprofilen für Klassen mit gleich vielen Clustern erfüllten nicht die unter 4.2.2 b) genannten Kriterien.

5. Interpretation und Schlußfolgerungen

Die Besprechung der Ergebnisse erfolgt zunächst nach den drei unter 2. genannten Fragestellungen.

(a) Der Tatsache, daß sich die beiden verwendeten Methoden in ihren Ergebnissen bezüglich des ersten Vergleichskriteriums, der ermittelten Gruppenzahl, stark unterscheiden, sollte nicht allzu großes Gewicht beigemessen werden. Die Zahl der Gruppen, in welche die Schüler einer Schulklasse partitioniert werden, ist bei der Clusteranalyse nur sehr unscharf bestimmbar. Und bei der Aufgabe der freien Gruppierung wird eine bestimmte Gruppenzahl von den Vpn häufig als willkürlich erlebt. Verschiedene Gruppen werden als zusammenlegbar, gebildete Gruppen nicht als homogen erlebt. Immerhin bleibt festzuhalten, daß der ermittelte Unterschied recht erheblich ist. Wie die verwendeten Koeffizienten zeigen, ist die Ähnlichkeit der Gruppenzusammensetzungen überzufällig groß. Es kann nicht davon die Rede sein, daß die Partitionierungen Ergebnisse von Zufallseinteilungen sind. Doch sind die Koeffizienten nicht in zufriedenstellender Höhe. Die beiden ver-

wendeten Methoden führen nicht zu vergleichbaren Ergebnissen. Etwa ein Drittel aller Schüler wurde im Durchschnitt unterschiedlich eingruppiert. Zwei mögliche Gründe für dieses Ergebnis sollen diskutiert werden.

Zunächst ist zu bedenken, daß beide Methoden unterschiedliche kognitive Leistungen von den Vpn verlangen. Bei der Methode der freien Gruppierung gilt es, Personengruppen in ihrer Gesamtheit zu ordnen. Welche kognitiven Prozesse dabei im einzelnen eine Rolle spielen, kann man nur vermuten. Nach der Theorie der Reizklassifikation (*Lilli 1975, S. 69f.*) ist ein Prozeß des Objektvergleichs anzunehmen. Ein Schüler wird hinsichtlich einiger dominanter Merkmale mit einem anderen Schüler verglichen. Schüler, die die gleichen Merkmale aufweisen, werden in die gleiche Kategorie eingeordnet. Das Wissen darum, daß man im Anschluß an die Gruppierung jede gebildete Gruppe mit einem treffenden Schlagwort zu bezeichnen hat, mag dazu geführt haben, eher mehr als weniger Gruppen zu bilden, um die anschließende Bezeichenbarkeit zu erleichtern. Eine andere Erklärungsmöglichkeit besteht darin, daß die Partitionierung von Schülern in sich ausschließende Gruppen keine geeignete Methode zur Erfassung der betreffenden Lehrerkognitionen darstellt. Anders ausgedrückt: Die Vorstellung einer Menge von Kategorien mit scharfen Grenzen bildet die Kategorisierung von Schülern durch Lehrer nicht angemessen ab. Ein Hinweis für diese Erklärung kann darin gesehen werden, daß viele Lehrer der freien Bildung von homogenen, abgeschlossenen Gruppen kritisch gegenüberstanden. Einigen fiel die Gruppierung sichtlich schwer. Die relativ häufig auftretende Mehrfachgruppierung bzw. Gruppenüberschneidung legt nahe anzunehmen, daß die Konkurrenz verschiedener simultan zu berücksichtigender Einteilungskriterien bei vielen Lehrern ein Gefühl des Unbehagens auslöste. Die relativ große Zahl von Lehrern, die trotz entsprechender Instruktion überlappende Gruppierungen bildeten, kann als Beleg dafür genommen werden, daß Lehrer Schüler eher als „fuzzy sets“ wahrnehmen, als Typen, zwischen denen fließende Übergänge bestehen. In solch einem Fall sind beide Techniken, die abgeschlossene, klar abgegrenzte Gruppen ähnlicher Objekte liefern, nicht geeignet, die zu erfassenden kognitiven Strukturen angemessen abzubilden. Neuerdings hat *Oldenburger (1986)* an den Daten von *Hofer (1981)* einen „clusterability“-Test durchgeführt. Er kommt zu dem Ergebnis, daß in den Daten wohl irgendwelche strukturellen Objekt- und Merkmalszusammenhänge bestehen, daß diese aber nicht clusteranalytisch repräsentierbar seien. Da die Reihenfolge in der Bearbeitung der beiden Aufgaben nicht permutiert worden war, ist schließlich auch ein Reihenfolgeeffekt nicht auszuschließen.

(b) Der Vergleich der Ergebnisse derselben Lehrer, bezogen auf zwei verschiedene Schulklassen, sollte über die intraindividuelle Stabilität von Kategorisierungen Auskunft geben. Das Design dieser Untersuchung ermöglicht es, den Vergleich mit zwei verschiedenen Methoden durchzuführen. Dies festigt die Glaubwürdigkeit der Ergebnisse. Ähnlichkeiten in den von gleichen Lehrern an verschiedenen Schulklassen gebildeten Gruppierungen lassen sich feststellen. Zwar lassen sich etwa zwei Drittel der Beschreibungsmerkmale für die bei der Methode der freien Gruppierung geschaffenen Gruppen denselben Merkmalen zuordnen, jedoch sind

auch erhebliche Unterschiede festzustellen, die der Erklärung bedürfen. Die Anwendung jeder Methode auf zwei Schulklassen kann als eine Art Paralleltestung angesehen werden. So kann ein Teil der Varianz auf die Unzuverlässigkeit der Meßinstrumente zurückgeführt werden. Darüber hinaus können Gruppierungen kontextspezifisch variieren. Denkbar ist, daß in Abhängigkeit von curricularen und pädagogischen Erfordernissen in verschiedenen Schulklassen unterschiedliche Gruppierungen resultieren. Auch können sich die Klassen z.B. hinsichtlich Klassenstufe, Unterrichtsfach, Klassengröße sowie in Merkmalen der in der Klasse zusammengefaßten Schüler unterschieden haben. Solche möglichen Einflußfaktoren konnten nicht systematisch konstant gehalten werden.

(c) Was den Vergleich der Ergebnisse verschiedener Lehrer in bezug auf dieselbe Schulklasse betrifft, so sind die Übereinstimmungen der Gruppierungen besonders niedrig. Die beiden berechneten Koeffizienten lagen zwar an der oberen Grenze des Zufallsbereichs, doch hoben sie sich nicht genügend davon ab, als daß von einer überzufälligen Übereinstimmung gesprochen werden könnte. Die gleichen Schüler wurden mithin von verschiedenen Lehrern ziemlich unterschiedlich gruppiert. Zunächst ist darauf hinzuweisen, daß die Lehrer zwar die gleichen Schüler, aber in unterschiedlichen Fächern unterrichteten, mit möglicherweise unterschiedlichen Methoden, auf unterschiedliche Ziele hin etc. Die unterschiedliche Perspektive, die aus solchen Kontextvariablen resultiert, mag für einen Teil der Varianz verantwortlich sein. Dazu kommt, daß dieselben Schüler sich bei verschiedenen Lehrern unterschiedlich präsentieren können, sei es in Abhängigkeit von der Person des Lehrers, sei es in Abhängigkeit von ihren individuellen Voraussetzungen (Interesse, Begabung etc.) für das speziell vom Lehrer vertretene Schulfach. Lehrer mögen auch unterschiedlich fähig und daran interessiert sein, ihre Schüler zu beobachten und zu angemessenen Einschätzungen über sie zu gelangen. Schließlich mag bei der Merkmaleinschätzung eine Rolle spielen, wie die Lehrer die vorgegebenen Eigenschaftsgegensätze interpretierten. Insgesamt ergibt sich ein Gegensatz zu den Untersuchungen von *Morine-Dershimer* (1978) und *Hofer* (1981), die interindividuelle Homogenität von Gruppierungen konstatieren konnten.

Eine Gesamtdiskussion der Ergebnisse hat abschließend danach zu fragen, wie die Qualität der verwendeten Methoden unter methodischem und unter theoretischem Aspekt zu beurteilen ist. Was das Ergebnis des Vergleichs zweier Methoden betrifft, so kann man sich der Forderung einiger Autoren (z.B. *Gigerenzer* und *Strube* 1978) nicht anschließen, bei der Erfassung von Kognitionen jenen Methoden den Vorzug zu geben, die Vpn möglichst wenig Vorgaben machen. Annahmefreie Methoden gibt es nicht. Die Aufgabe der freien Gruppierung lieferte weder zuverlässigere Werte, noch kamen die Vpn damit besser als bei der Eigenschaftsbeurteilung zurecht. Beide Methoden basieren auf unterschiedlichen Annahmen, fordern den Vpn unterschiedliche Leistungen ab und erfassen z.T. unterschiedliche Aspekte. Trotz der Konstatierung eines Gutteils an Übereinstimmung hat die Untersuchung unter theoretischem Aspekt Zweifel an der Berechtigung des beiden Verfahren zugrundeliegenden Modells aufkommen lassen. Das Modell, daß die Vorstellungen von Lehrern über Schüler als Kategoriensystem beschreibbar sind, enthält Annah-

men, von denen jede einzelne, mehrere oder alle unzutreffend sein können: die Annahme, daß das Kategoriensystem erschöpfend ist, daß die Kategorien disjunkt und daß Objekte innerhalb der Kategorien homogen und zwischen ihnen heterogen sind. Möglicherweise sind beide Methoden gleichermaßen ungeeignet, kognitive Schülerrepräsentationen angemessen zu erfassen. Oldenburger (1986) stellte in seiner Reanalyse der Daten aus Hofer (1981) fest, daß die Beurteilungen der Lehrer nicht wie Punktwolken im euklidischen Raum organisiert sind. Offen sei, welche Art von Struktur diese wechselseitig verbundenen Netze aufweisen und welches mathematische Modell zu ihrer Abbildung herangezogen werden könne. Vielleicht sind für Lehrer nicht so sehr Persönlichkeitseigenschaften auffällig, sondern individuelle Verhaltensweisen, die in Verhaltenssequenzen eingebettet sind. Gespeichert werden Verhaltensskripts und weniger Merkmalskombinationen. Bestimmte Schüler können in verschiedenen typischen Skripts vorkommen, andere in keinem. Die vorliegende Arbeit kann die Frage nach alternativen Erhebungs- und Verrechnungsverfahren nicht beantworten.

Literatur

- Bergs, S.: Optimalität bei Clusteranalysen. Experimente zur Bewertung numerischer Klassifikationsverfahren. Unveröffentlichte Dissertation. Münster 1980.
- Brophy, J. E., Good, T. L.: Teacher-student relationships, Causes and consequences. Holt, Rinehart & Winston, New York 1974.
- Cantor, N., Mischel, W.: Prototypes in person perception. In: Berkowitz, L. (Hrsg.): Advances in Experimental Social Psychology (Bd. 12). Academic Press, New York 1979, 3–53.
- Cunningham, W. G.: The impact of student-teacher pairings on teacher effectiveness. American Educational Research Journal (1975) 12, 169–189.
- Garner, J., Bing, M.: Inequalities of teacher-pupil contacts. British Journal of Educational Psychology (1973) 43, 234–243.
- Gigerenzer, G., Strube, G.: Zur Revision der üblichen Anwendung dimensionsanalytischer Verfahren. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie (1978) 10, 75–86.
- Hofer, M.: Schülergruppierungen in Urteil und Verhalten des Lehrers. In: Hofer, M. (Hrsg.): Informationsverarbeitung und Entscheidungsverhalten von Lehrern. Beiträge zu einer Handlungstheorie des Unterrichtens. Urban & Schwarzenberg, München 1981, 192–221.
- Hofer, M. (Hrsg.): Sozialpsychologie erzieherischen Handelns. Wie das Denken und Verhalten von Lehrern organisiert ist. Hogrefe, Göttingen 1986.
- Lilli, W.: Soziale Akzentuierung. Kohlhammer, Stuttgart 1975.
- McFarland, W.: Are girls really smarter? Elementary School Journal (1969) 70, 14–19.
- Morine-Dersheimer, G.: Teacher conceptions of pupils. Vortrag gehalten auf dem Kongreß der American Educational Research Association, Toronto 1978.
- Oldenburger, H. A.: Does a tendency to group pupils or attributes exist in teachers' cognitions/judgments? In: Ben-Peretz, M., Bromme, R., Halkes, R. (Hrsg.): Advances of Research on Teacher Thinking. Swets & Zeitlinger, Lisse 1986, 186–200.
- Rosch, E.: Principles of Categorization. In: Rosch, E., Lloyd, B. B. (Hrsg.): Cognition and Categorization. Erlbaum, Hillsdale 1978, 27–48.
- Silberman, M.: Behavioral expression of teacher's attitudes toward elementary school students. Journal of Educational Psychology (1969) 60, 402–407.
- Thelen, H. A.: Classroom grouping for teachability. Wiley, New York 1967.
- Ward, J. H.: Hierarchical grouping to optimize an objective function. Journal of the American Statistics Association, (1963) 58, 236–244.
- Wishart, D.: An algorithm for hierarchical classification. Biometrics (1969) 25, 165–170.

Verfasser:

Prof. Dr. Manfred Hofer, Dipl.-Hdl. Bernd Köpke, Lehrstuhl für Erziehungswissenschaft II, Universität Mannheim, Schloß, 6800 Mannheim 1